



Red Latinoamericana de Formadores en Fact-checking



WEBINAR

Verificar lo inédito: nuevos retos y herramientas frente a la IAG en el periodismo



Jhoandry Antonio
Suárez Urdaneta



Este webinar propone un espacio de reflexión y actualización sobre los desafíos emergentes que estas tecnologías plantean para la verificación de contenidos y explora recursos concretos que pueden integrarse en la formación periodística universitaria.



Miércoles 18 de junio



12 hs



11 hs



10 hs



9 hs

Organizan



chequeado



COLOMBIACHECK

No comunique

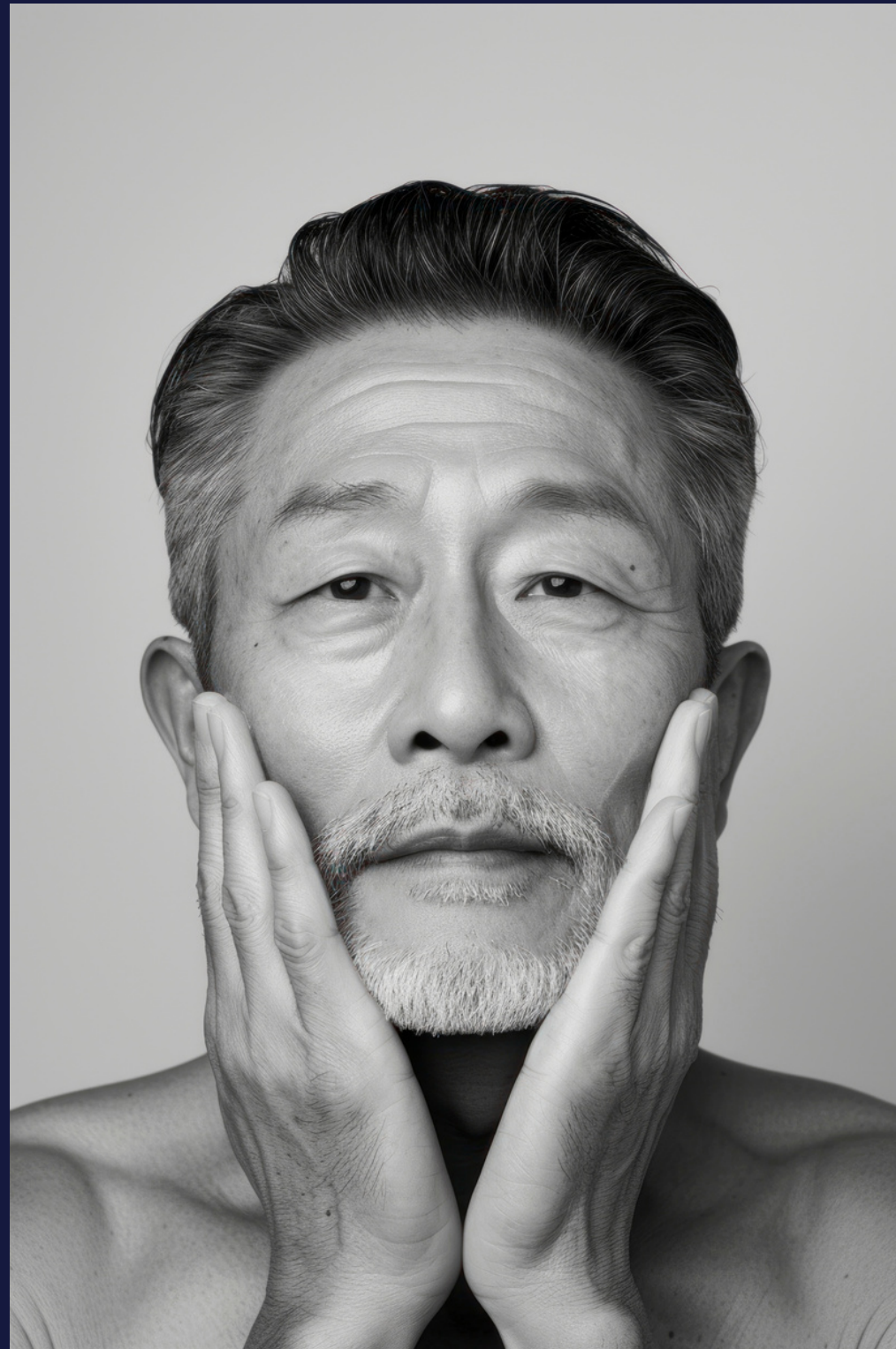


Financiado por
la Unión Europea

¿IMAGEN REAL O CREADA CON IA?



¿IMAGEN REAL O CREADA CON IA?



¿CÓMO DEFINIRÍAS LA INTELIGENCIA ARTIFICIAL? EN UNA PALABRA



<https://www.menti.com/alhx176aiigd>

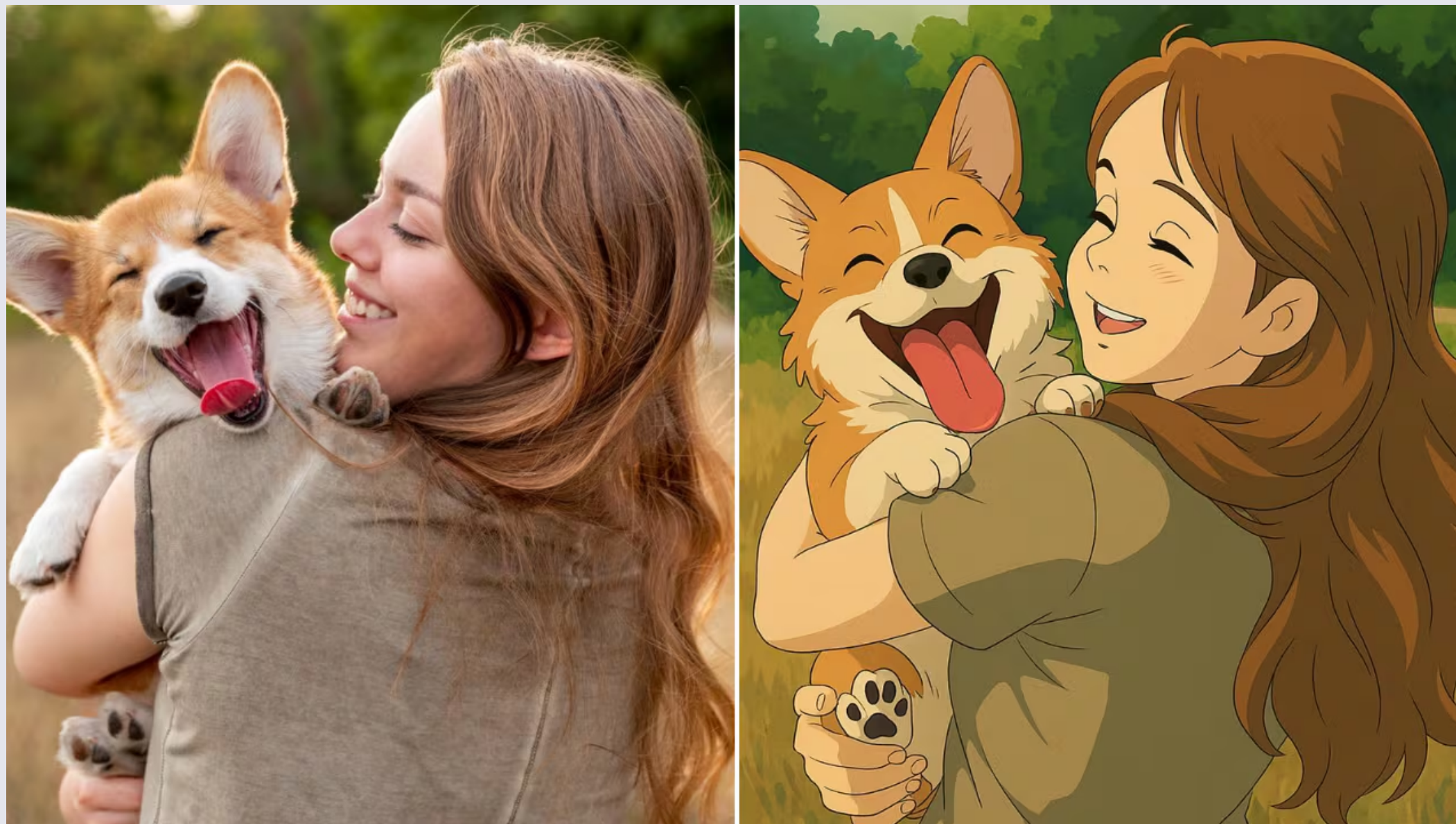
HABLEMOS DE LA IA

“La inteligencia artificial (IA) es un **conjunto de tecnologías** que permite a las máquinas simular capacidades propias de la inteligencia humana, como la resolución de problemas, el aprendizaje y la ejecución de tareas de manera autónoma”

Cole Stryker y Eda Kavlakoglu

PERO ENTONCES, ¿QUÉ ES LA IAG?

Es lo que hace posible esto...



“La IA generativa, a veces denominada ‘Gen AI’, hace referencia a modelos de deep learning que pueden **crear contenidos** originales complejos, como texto extenso, imágenes de alta calidad, video o audio realistas y mucho más, **en respuesta a una solicitud o petición del usuario**”.

Cole Stryker y Eda Kavlakoglu

EN RESUMEN...

**“La inteligencia artificial
no es una herramienta,
es un agente que puede
crear por sí mismo”.**

Yuval Noah Harari,
antropólogo y escritor.

LA IA USADA PARA DESINFORMAR

La masificación y gratuidad de herramientas de IAG permiten que se creen contenidos falsos o engañosos con un impresionante realismo. Esto facilita la diseminación de desinformación en redes que favorecen las publicaciones con mejor engagement, sean verdaderos o no.



HERRAMIENTAS PARA DETECTAR VIDEOS O IMÁGENES CREADAS CON INTELIGENCIA ARTIFICIAL.

¿CÓMO FUNCIONAN ESTAS HERRAMIENTAS?

- Utilizan **algoritmos entrenados** con diversos conjuntos de datos para identificar las anomalías sutiles que pueden indicar una falla profunda.
- Estudian **patrones** en miles de videos, aprendiendo a discernir lo auténtico de lo manipulado.
- Pueden **detectar detalles** que se pueden deslizar al ojo humano, como ligeras distorsiones en el fondo o cambios faciales diminutos.
- También son capaces de analizar **los metadatos** sobre el origen del contenido si fue creado con una IAG.



Semana



"LA FOTO QUE LO INCRIMIN



GUSTAVO PETRO DICE NO CONO
A PAPA PITUFO PERO LAS FOTO
DICEN LO CONTRARIO SERÁ Q
PETRO SUFRE DE AMNESIA O QU
GENERAR DISTRACCIÓN.



EL PANORAMA DE ESTAS HERRAMIENTAS

- La mayoría de herramientas han sido desarrolladas por empresas que tienen servicios de IA.
- No existe tanta variedad de detectores gratuitos disponibles.
- Muchas aplicaciones comienzan siendo gratuitas y luego las restringen a pagas.
- No son infalibles.

INVID DEEPFAKE

hivemoderation.com/ai-generated-content-detection

- Para videos.
- Está en fase experimental.
- Requiere registro.

 **Deepfake**

Utiliza un clasificador de aprendizaje automático que devuelve la probabilidad de que un video contenga rostros manipulados por IA (intercambio de rostros y recreación facial).

 Advertencia: esta es una herramienta experimental en desarrollo. Ha sido mejorada dentro del proyecto veraai.eu, pero aún puede tener errores o fallos. Úsala con precaución. Siempre verifica si existe un video anterior (con fotogramas clave). En caso de duda, intenta usar otras herramientas de terceros para confirmar cualquier resultado. Se han reportado algunos casos de falsos positivos, especialmente en videos con multitudes y rostros borrosos en el fondo de las imágenes.



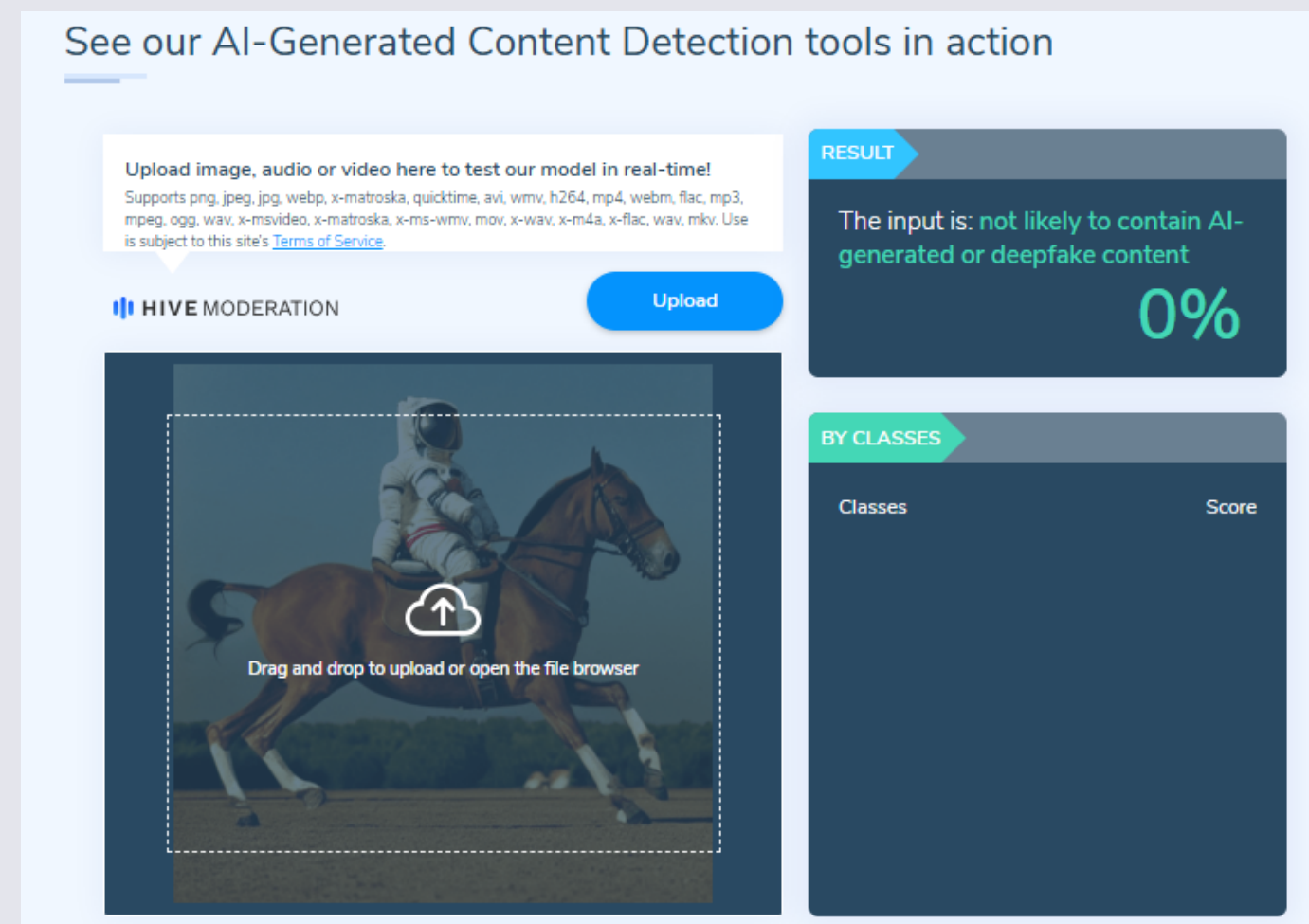
 Fichero local



HIVE MODERATION

hivemoderation.com/ai-generated-content-detection

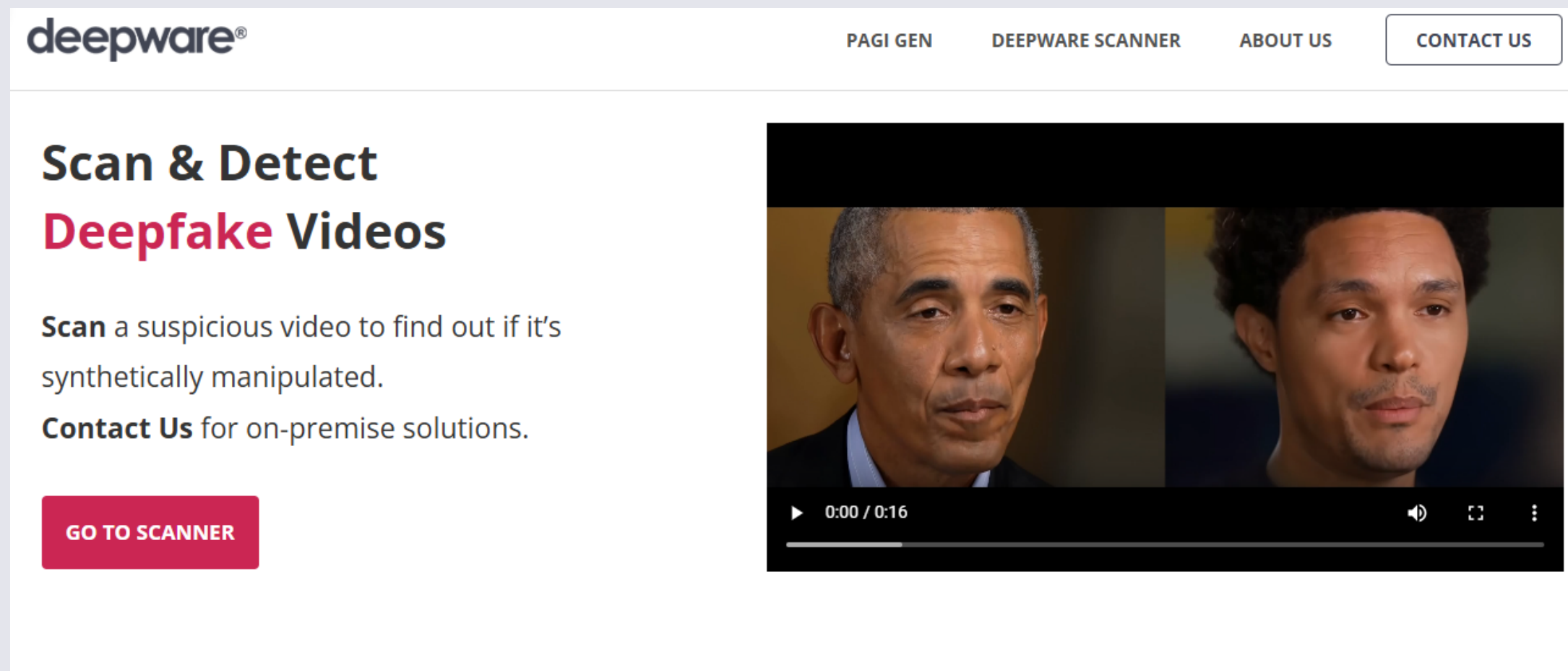
- Para videos, imágenes y textos.
- Dispone de un plugin para el navegador.
- No requiere registro.



DEEPWARE

deepware.ai/

- Funciona para videos.
- Es más preciso con videos de personas.
- No requiere registro.



The screenshot displays the Deepware website's main interface. At the top, the 'deepware®' logo is on the left, and navigation links for 'PAGI GEN', 'DEEPWARE SCANNER', 'ABOUT US', and a 'CONTACT US' button are on the right. The main content area features the heading 'Scan & Detect Deepfake Videos', where 'Deepfake' is highlighted in red. Below this, a description states: 'Scan a suspicious video to find out if it's synthetically manipulated.' and 'Contact Us for on-premise solutions.' A red button labeled 'GO TO SCANNER' is positioned below the text. To the right of the text is a video player showing a side-by-side comparison of Barack Obama's face, with a progress bar at the bottom indicating '0:00 / 0:16'.

SIGHT ENGINE

<https://sightengine.com/detect-ai-generated-images>

- Solo para imágenes.
- Puede indicar la IA que creó la imagen.
- No requiere de registro.

Detect AI-generated images at scale

Our AI image detector automatically detects images from popular AI generators. No watermarks needed.



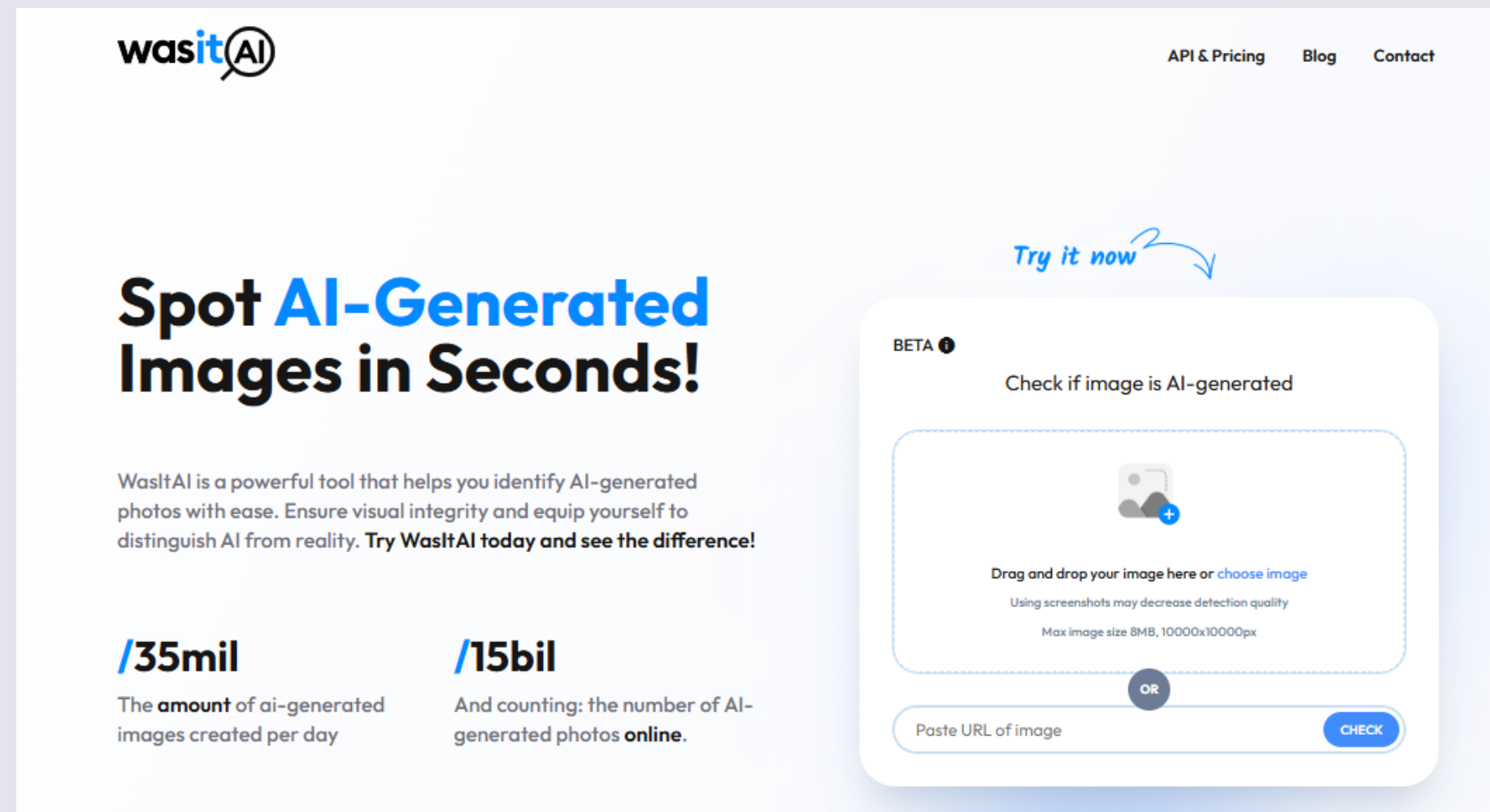
```
1  "status": "success",
2  "request": {
3    "id": "req_fPD38fsQgC1E2Ed8v",
4    "timestamp": 1709634469.399,
5    "operations": 1
6  },
7  "type": {
8    "ai_generated": 99%
9  },
10 "media": {
11   "id": "med_fPD3iQzaZR2E193Oxb",
12   "uri": "image.jpg"
13 }
```

[CHECK IMAGE](#)

WASITAI

<https://wasitai.com/>

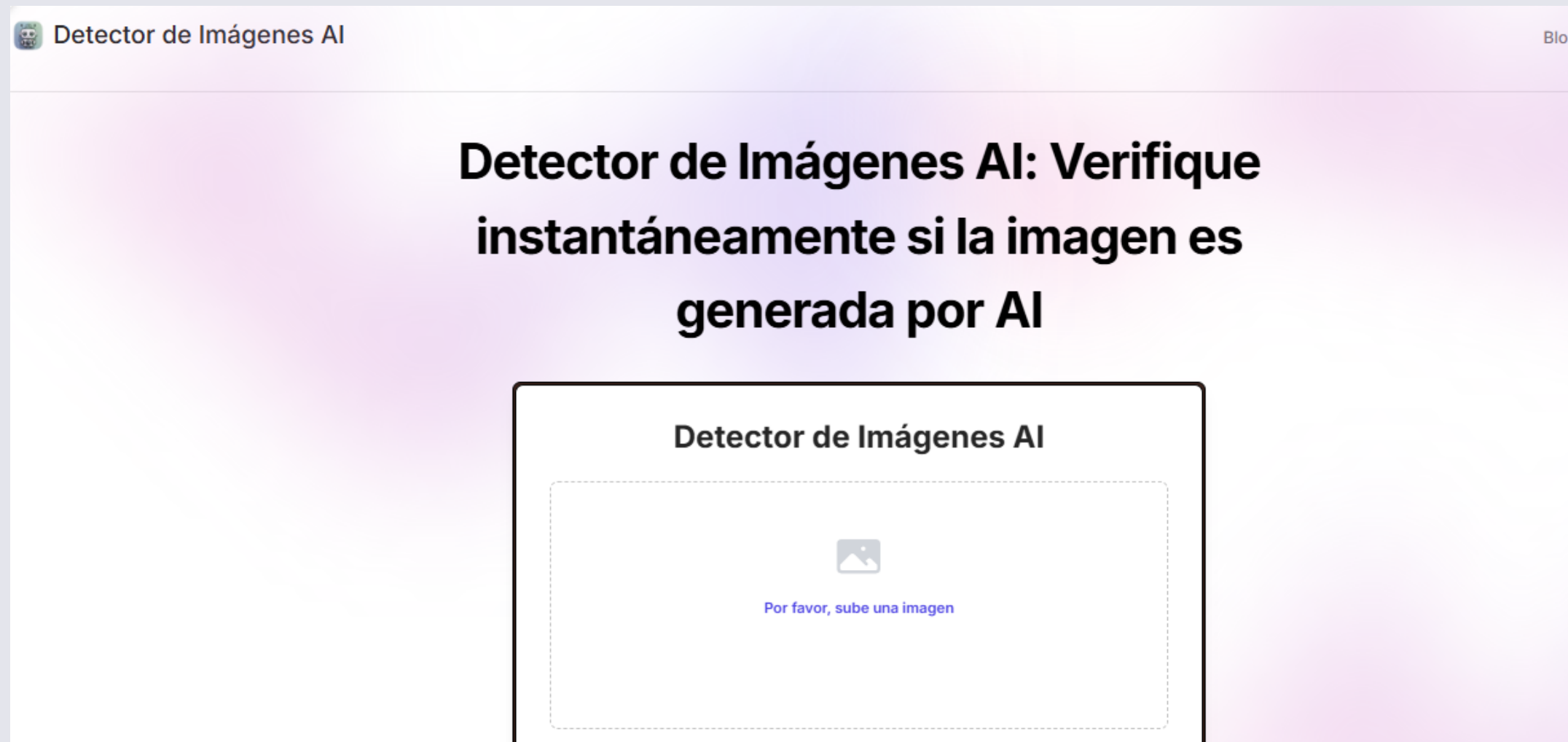
- Solo para imágenes.
- No ofrece un resultado detallado.
- No requiere de registro.



DETECTOR DE IMÁGENES AI

aiimagedetector.org/es

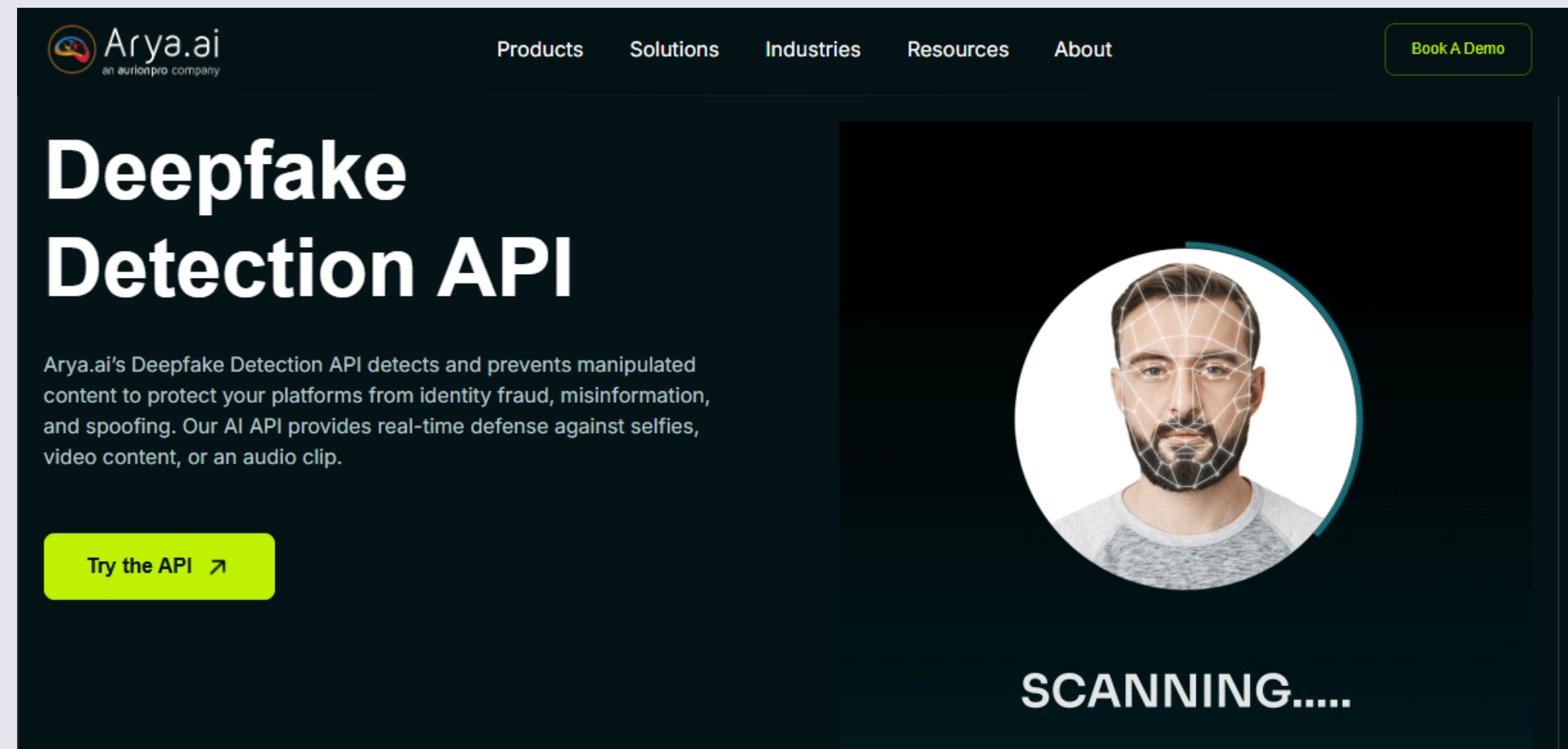
- Solo para imágenes.
- No ofrece un resultado detallado.
- No requiere de registro.



ARYA

arya.ai/apex-apis/deepfake-detection-api

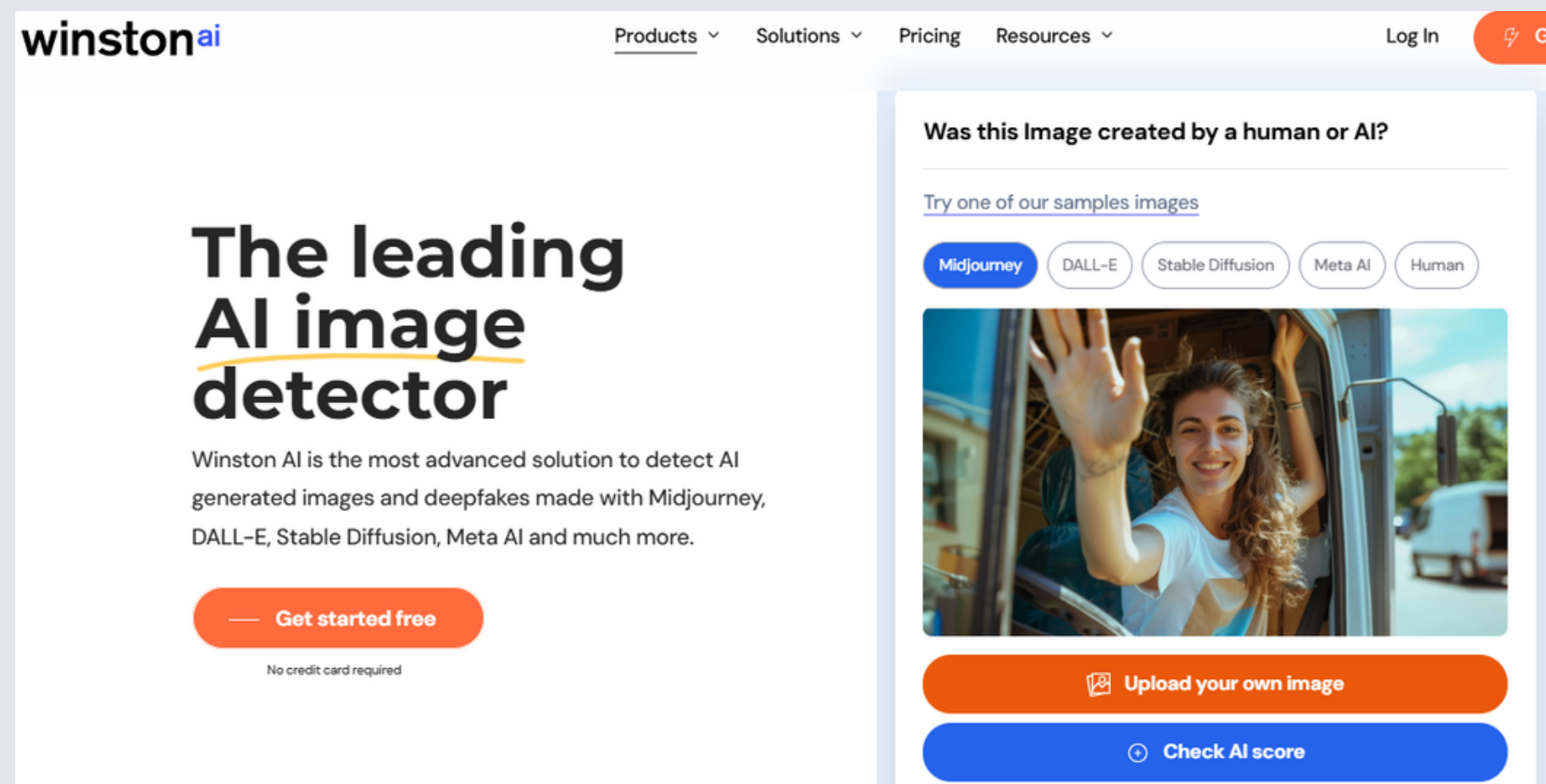
- Funciona para videos, imágenes y audios.
- Permite un número limitado de detecciones gratis.
- Requiere registro.



GO WINSTON

gowinston.ai/ai-image-detector/

- Solo para imágenes.
- Permite un número limitado de detecciones gratis.
- Requiere registro.



PROBEMOS LAS HERRAMIENTAS



<https://drive.google.com/file/d/1cUFmmWC1b1Elug6A9jQz4WS8OfPUr4Sr/view?usp=sharing>



https://drive.google.com/file/d/1M8_L7_e4qL-37o4cXOxujXTL7lqhHKoT/view?usp=sharing



CONSEJOS PARA USAR ESTOS DETECTORES.

01

Procura siempre subir el video o la imagen en la mejor calidad posible. Evita los pantallazos o capturas.

02

Siempre prueba con varios detectores. No te quedes solo con uno, pues todo se basa en probabilística.

03

Expón a tu audiencia los resultados que te dan los detectores, así no coincidan en su conclusión.

CONSEJOS ADICIONALES

- Observa detenidamente los detalles: la proporción de las partes del cuerpo, errores en los dedos, texturas borrosas, las sombras.
- Evalúa si los elementos del fondo son lógicos y consistentes con la persona o el objeto principal.
- Si se trata de un video, extrae fragmentos y analiza si hay elementos extraños.
- Detalla movimientos raros en la boca y si el audio va en sincronía con la voz.
- Revisa marcas de agua, a veces están en las esquinas de la imagen.
- Recurre a la búsqueda inversa en Google.



**¿DUDAS,
PREGUNTAS,
COMENTARIOS?**

GRACIAS



jhoandry.suarez@colombiacheck.com